DOCUMENT RESUME

ED 277 225 FL 016 154

AUTHOR Svartvik, Jan
TITLE The Segmentation of Impromptu Speech.
PUB DATE 82
NOTE 17p.; In: Enkvist, Nils Erik, Ed. Impromptu Speech: A Symposium. Papers Contributed to a Symposium on Problems in the Linguistic Study of Impromptu Speech (Abo, Finland, November 20-22, 1981); see FL 016 148.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Classification; *Computer Oriented Programs; *Computer Software; Dialogs (Language); *Discourse Analysis; Language Research; Linguistic Theory; Oral Language; Phrase Structure; *Research Methodology; *Speech Acts; Spontaneous Behavior; Structural Analysis (Linguistics); Syntax
IDENTIFIERS *Spontaneous Speech

ABSTRACT

A computer program for classifying elements of a language corpus for large-scale analysis is discussed. The approach is based on the assumption that there is a natural unit in speech processing and production, called a tone unit. The program "tags" the five grammatical phrase types (verb, adverb, adjective, noun, and prepositional) to provide a basic syntactic description of each tone unit of a text. In the first part of the program, all complete words are tagged for word class. In the second part of the program, all word-class tags are scanned by tone unit, and labels assigned to grammatical phrases. Third, the tone units are analyzed from right to left. A demonstration of the program design analyzes some verb phrases. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED277225

THE SEGMENTATION OF IMPROMPTU SPEECH

Jan Svartvik
University of Lund

FL 016 154

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

ENKvist

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

Jan Svartvik
University of Lund

The segmentation of impromptu speech

1 Why is speech segmentation a problem?

The segmentation problem in spoken discourse is particularly acute on the word level: 'In fluent speech there are no physical cues that consistently and reliably indicate where one word ends and the next begins'; '. . . less than 40% of all word boundaries are marked by some physical event' (Cole & Jakimik 1980:138). Since the material we are working with at the Survey of Spoken English, the London-Lund Corpus, has been compiled mainly for grammatical, prosodic and pragmatic studies, the transcription uses institutionalized, orthographic words (except for specific items, word partials, etc.).[1] Fortunately, we can therefore leave the word segmentation problem to our fellow students of speech recognition.

Our segmentation problem is to be found on the syntactic level where the traditional unit 'sentence' is often impossible or, at best, awkward to work with. Crystal has given three reasons why sentence identification and classification is problematic in spoken material: indeterminate connectivity, indeterminate ellipsis and intercalation of structures (1980:155). Clauses and phrases are easier to identify in speech. But one problem here is that we have no grammatical model based on clauses and phrases; another, and a more serious one, is that 'identification' does not seem to be an ambitious enough goal. I would like to take 'segmentation' to mean something like 'a mode of analysis that satisfies reasonable demands for a linguistically adequate model and also reflects the natural units of discourse production, processing and perception'.

This may be too tall an order, but I really see no alternative for the reason that speech is produced, processed and perceived under conditions which are radically different from those of writing and thus bound to result in a different linguistic organisation of information. Hence 'the sentence may not be a practical or even "proper" unit for understanding speech processes' (Weiner & Goodenough 1977:213). So, before discussing possible candidates

for a proper unit of speech segmentation, we will consider under what conditions the speech process takes place.

2 A classification of speech events

A typical feature of most spoken discourse is that it permits less planning than written. But speech varies in this and other respects depending on the speech situation. We have classified the texts of the London-Lund Corpus according to five dimensions: dialogue, face-to-face, private, surreptitious recording and radio broadcast. As can be seen in the matrix representation of Figure 1, these five features produce five text groups with eight subgroups. By and large, these groups can be seen as forming a scale from least planned (face-to-face conversation) to most planned speech (public, prepared oration). A classification in terms of the amount of planning that characterizes a speech event has a bearing on the segmentation problem insofar as it is likely to become more difficult the closer we move towards the most-typical--impromptu speech end of the continuum.[2]

The kind of speech that will be discussed here is the most common and typical one, ie informal, impromptu dialogue, as exemplified in a prosodic transcription in Figure 2.

3 How is dialogue produced?

Speakers have to do two things as they talk: first, to determine what to say and what linguistic form to use; second, to deliver these linguistic goods. In dialogue further demands are made on the participants: adhering to unwritten turn-taking rules, processing current speaker's contribution while planning their own, deciding how to say it and then actually say it, attracting the listener's attention and maintaining it by producing comprehensible and worthwhile-listening-to output, eliciting response, and so forth. The structure of discourse is said to be of two kinds (Clark & Clark 1977:236): hierarchical, which is the outgrowth of the speaker's purpose of talking; and local, which is the necessarily serial relation of contributions to a conversation, ie planning ahead is really possible only after the last contribution. This is the reason why

Figure 1. Text classification of the London-Lund Corpus

Text groups		Subgroups	Dialogue	Face-to-face	Private	Surreptitious	Radio	Text labels	Sub-group total	Text group total
Abbre-viation	Broad description									
CON	Face-to-face conversation	A	+	+	+	+	–	S.1.1-14, S.2.1-14, S.3.1-6	34	46
		B	+	+	+	–	–	S.4.1-7, S.5.8-11, S.6.2[1]	12	
TEL	Telephone conversation	C	+	–	+	+	–	S7.1-3, S.8.1-4, S.9.1-3[2]	10	10
DIS	Discussion, inter-view, debate	D	+	+	–	–	+	S.5.1-7, S.6.1, S.6.3-5[3], S.10.8[4]	12	12
PUB	Public, unpre-pared commen-tary, demon-stration, oration	E	+	+	–	–	–	S.11.1, S.11.4-5	3	12
		F	–	+	–	–	+	S.6.6[5], S.11.3	2	
		G	–	–	–	–	+	S.10.1-7[6]	7	
PRE	Public, prepared oration	H	–	+	–	–	–	S11.2, S.12.1-6[7]	7	7
			Number of "+" texts:						87	87
			71	70	56	44	21			

NOTES TO FIGURE 1

1 Subgroup B is composite; individual speakers can be marked for surreptitiousness.

2 S.9.3 is ansaphone calls.

3 Subtext S.6.4b, which constitutes only a minor part of Text S.6.4, is not a radio recording but a discussion between an ex-nurse and a research worker on a topic similar to that of S.6.4a.

4 Subtext S.10.8b is not a radio recording. It has been placed in S.10.8 because it is very short, less than 100 tone units. Furthermore, the speaker in S.10.8b is one of the speakers in S.10.8a and both subtexts deal with popular science.

5 S.6.6 is an interview which, as it turned out, virtually became a monologue.

6 Subtext S.10.6a is a TV recording. Subtext S.10.7c is not a radio recording (like the main part of Text S.10.7) but a demonstration.

7 There are two speakers in S.12.3. However, it has been placed in this text group since it is virtually a monologue.

Figure 2: A text sample from the London-Lund Corpus (S.1.1)

A 222 ⁎we ‖shouldn't⁎

>B 221 from YŎU‖

A 223 and we ‖shouldn't be im▵posing on you in VAΔCÀTION ⁎ ▷time‖⁎ 224 «there's ‖no» ΔPŌINT‖

B 225 ⁎«well ‖as its»⁎

>A 226 ‖you SĒE‖ 227 of of ‖asking for ▵this particular DÀTE «the» {‖first of MÀY‖} ‖ 228 «was [s] that» [ə:m] our ex‖ternal EXĀMINER‖ · 229 ‖as well as OUR▵SĒLVES of ▷course‖ ⁎ · ⁎ 230 [ə:] ‖would be Δfree during TÈRM-▷TIME‖ + - +

B 231 ⁎‖[m̄]‖⁎ 232 + ‖[mhm̀]‖+

>A 233 be ‖freed ˇby the time term's ÈNDED‖ - 234 [ə:m] you're ‖very kind old SÁM‖ - - 235 ˇBLÈSS ▷you‖ 236 well ‖that finishes THÀT‖ · 237 [ə:m] · «now» ‖what was the ŎTHER thing {I ‖wanted to ▵ĀSK ▷you‖} ‖ · 238 ‖[i] ˇis · ‖is it ˇthis YÈAR‖ 239 that [ə:] ‖NĪGHTINGALE ▷goes‖ - -

B 240 [ə:] no ‖NÈXT ▷year‖ - -

A 241 [ə:m] · ‖sixty- [f] ⁎-⁎ four SIXTY-FĪVE‖ + · +

B 242 ⁎‖SIXTY-FĪVE‖⁎ 243 + ‖YÈAH‖+

>A 244 «I thought it was ‖BEFŌRE» sixty-▷five‖ 245 so it's ‖not until - Δnext year that ⁎the job «will be» ▵ĀDVERTISED‖⁎

B 246 ⁎‖January I sup▵pose there⁎ ▵may be an ▵interview round about JĀNUARY‖

A 247 ‖YÈAH‖ - 248 [ə:m] ‖YŎU heard anything {a‖bout THÍS‖} ‖ ·

B 249 ˇnothing at ΔĀLL «YÉT»‖ - -

A 250 [ə:m] - - you've ‖not heard ΔPĒEL ▷mentioned {in ‖this CONNĒXION‖} ‖

B 251 well ‖NĪGHTINGALE ▷mentioned it {CĀSUALLY to ▷me‖} ‖ 252 that ‖Peel ΔMĪGHT try FŌR it‖ 253 do you know anything ‖definite «ABŌUT him‖»

A 254 ‖NŌ‖ - 255 [ɔ:] but ‖[sʌmbp ə] it ‖may have come from the same ΔSŎURCE a▷gain‖

B 256 ‖well ΔI mentioned it to YŎU‖ 257 if you REI‖MÉMBER‖ 258 ⁎«1 to 2 sylls»⁎

A 259 ⁎‖yes I⁎ ΔKNŌW‖ · 260 I know ‖YŎU did‖

B 261 ‖[mhm̀]‖

A 262 [ə:m] I «[ri'ʒmʌ] at ‖least» NŎW‖ 263 I [ə] you ‖come to MĚNTION it‖ 264 I REI‖MÈMBER that you ▷did‖ - 265 [ə] but I ‖heard it · [ə] mentioned by somebody ÊLSE‖ - 266 I ‖think ΔWĀTT‖ - 267 I'm ‖not SŪRE‖ ⁎-⁎ 268 [ə:m] - - ‖and [ə:m]

B 269 ⁎[m]‖[hm̀]‖⁎

it is so difficult to carry on a relaxed conversation in a foreign language we do not know well: the privilege of talking time-out for planning your own contribution, which exists in prepared forms of speech and in writing, is just not available in spoken interactive discourse. Since this paper deals with segmentation, it is the local kind of discourse structure that concerns us here, in particular the question in what linguistic form discourse is delivered by the speaker and processed by the listener.

In the psycholinguistic literature it is generally assumed that the local planning unit it is the sentence but that 'the words within a sentence . . . are not planned at once. Speakers appear to work from a skeleton plan of a sentence and select the words roughly constituent by constituent' (Clark & Clark 1977:257). Most of what we know about the execution of speech plans is based on the analysis of speech errors and grammatically incoherent utterances. What these tell us, among other things, is that planning and execution alternate since the speaker cannot plan very far ahead. The speaker 'has a relatively complete, though perhaps sketchy, cognitive structure in mind, but . . . a syntactic form for the whole of this structure has not been formulated berofe the utterance begins' (K. Brown 1980:31).

In some of these areas considerable progress has been made by psycholinguists, who are particularly good at conducting experiments, measuring pauses in microseconds, etc., but often employ crude linguistic analyses and fail to make use of authentic, prosodically analysed utterances.[3]

4 Is there a natural unit of segmentation?

Phoneticians have noted that intonation patterns are structured in terms of the speaker's breathing patterns and recognize the 'breath group' as 'one of the basic, primitive phonetic features that must be present in all languages' (Lieberman 1980:192). Typical features of speech units such as limited variation in length, as compared with units in writing, and patterns of falling fundamental frequency are thus linked with respiratory activity. The breath group is 'a signalling unit that has clear articulatory and acoustic correlates that reflect the presence of neutral mechanisms that govern its production and perception' (Lieberman 1980:198).

Two acoustic features seem to be of special intresse to us here: first, the tendency for the fundamental frequency to drift downward within the breath group at a rate that varies inversely with length, ie the declination is relatively independent of the length of the group; second, final lengthening, ie 'the time per word increases with the length of an utterance and also tends to increase with serial position within the utterance' (Sternberg *et al.* 1980: 518ff.).[4] These features can be taken to provide evidence of cohesion within the prosodic unit and of advance planning.

What phoneticians call 'breath group' has been given many other names, eg 'information unit' by Halliday, 'idea unit' by Chafe, indicating that there is a correlation between the unit of production and the structure of concepts: 'one aspect of the organization of discourse is the distribution of information into sequential information units' (Kroch & Hindle 1981:28); McNeill talks about 'meaning units produced as single outputs' (1979:241); Chafe suggests that idea units are 'linguistic expression of focuses of consciousness' (1980: 15).

5 What is the grammatical structure of the tone unit?

There is, then, considerable support from several quarters for the existence of a natural unit in speech processing and production. We will assume that it corresponds to the 'tone unit'.[5] If the tone unit is a natural, conceptual unit, it should be of interest to know how it is grammatically structured. A close study of authentic speech material will reveal that there is no general, consistent correspondence between grammatical and prosodic units.[6]

In a study of Survey texts (Quirk *et al.* 1964), the tone unit was found to have an average length of 5.3 words, with the highest number of occurrences clustering between 3 and 5 words; 61 per cent of the units had lengths fairly evenly distributed between 2 and 7 words. Chafe's idea units have a mean length of about 2 seconds, or about 6 words each (1980:14). In her contribution to this symposium, Déjean Le Féal states that she found the length of segments consisting of 'phrase sections located between two pauses' to be from 1 to 9 words, – and most of them less than 7 words long – in impromptu speech (compared with segments from 3 to 22 words long in texts read aloud).

Quirk *et al.* found there was considerable correlation between the length of the tone unit and their grammatical content. The major correspondences were found to be as follows (capital letters indicate words carrying nuclei and '//' indicates end of the unit).

Two clauses (8 words):
I really think that he will come TOMORROW//
One clause (5-6 words):
he will arrive by train TOMORROW//
Verbal group plus adverbial (4-5 words):
will arrive at ten THIRTY//
Nominal group as subject (3-4 words):
the train from LONDON//
Prepositional phrase as adverbial (3-4 words):
at this PLATFORM//
Miscellaneous nominal groups (2-3 words):
and the SOUTH//
Adverb (1 word):
OBVIOUSLY//

The overall impression is that there is indeed correlation between tone units and grammatical units but that it cannot be stated in a simple, general way in terms of constituents, phrases, clauses or sentences. Two examples: we found there was a tendency for inter-unit junctions after clauses operating as post-modifiers in nominal groups:

the man I know// is over there//

and within clauses operating as complements in another clause:

I think this man// will do it//

Although such junctions can no doubt be explained in linguistic terms, we do not at present have enough knowledge of this area to make predictive statements about grammatical/prosodic distributions in general.

The study of correspondences between grammatical and prosodic features reported in Quirk *et al.* (1964) was carried out during the early years of the Survey of English Usage. It was an experiment based on a limited material. We are now in a better position to continue this research: the corpus totals some 435,000 words, it is available on magnetic computer tape and computer facilities are generally much better than in the early sixties. At the Survey of Spoken English we are currently engaged in 'tagging' the corpus grammatically so as to enable us to conduct large-scale studies with the help of a computer.

6 The design of a tagging program

Our tagging is carried out with a computer program that is semiautomatic, interactive and heuristic. The procedure can be described briefly as follows. (For a more detailed and technical description see Svartvik & Eeg-Olofsson, forthcoming.)

In the first part of the tagging program, words are tagged for word-class by a semiautomatic tagging program that assigns word-class labels to all complete words in a text. The current rate of correct automatic tagging produced by this program is about 80 per cent. We make use of a small lexicon containing the 1000 most frequent words in the text at the actual point in the analysis, ie the word types in the lexicon will vary during the analysis and, in the case of multiple tags, the statistically most probable tag will be chosen. In addition to the lexicon, we have a set of suffix and context frame rules. The suffix rules will, for example, tag a word in *-ly* which is not in the lexicon as 'adverb'. The context frame rules are used for the disambiguation of frequent items. Thus *to* will be tagged as 'infinitive' marker when an infinitive follows, otherwise as 'preposition'. The accuracy can no doubt be increasing by adding more and better context frame rules based also on other criteria, such as prosody and position in the tone unit.

The second part of the program is a syntactic phrase parser. All word-class tags are scanned, tone unit by tone unit, and labels assigned to grammatical phrases. There are two features in the design of this program which may be of more general linguistic interest. First, the algorithms are based on word-class tags, not lexical words. Second, the phrase recognition rules are recursively applied in the following order:

(VPH) Verb phrases
(APH) Adverb phrases
(JPH) Adjective phrases
(NPH) Noun phrases
(PPH) Prepositional phrases

Third, tone units are analysed from right to left. The decision to use this search order was taken chiefly because it seems easier to identify and isolate phrases in this way as English phrase structure typically has the head to the right and, furthermore, the phrase heads often carry the nucleus. Although this order was adopted for technical reasons rather than in an attempt to simulate human processing, it is of interest to note that psycholinguists are in fact divided in respect of processing direction in human speech. Cf, for example, the statement by Cooper & Paccia-Cooper: 'It is not necessary that words be considered in left-to-right fashion at all computational stages involved in the planning of speech. For example, at an early stage of coding, the speaker may decide on a particular noun phrase without regard to its linear position in the speech output . . .; the experiments do not provide a satisfactory answer to the question of whether left-to-right processing accompanies syntactic stages in *ongoing* speech production. In short, while the phonetic output of speech proceeds in left-to-right fashion by definition, there is still no assurance that left-to-right processing accompanies all of the syntactic stages of coding which occur internally' (1980:198-199).

7 *An algorithm for verb phrase tagging*

The phrase tagging program is designed to provide a basic syntactic description of each tone unit of a text, in the first instance in terms of the five grammatical phrases. For a demonstration of the program design and a description of the technique for parsing one such grammatical unit, this section will outline our analysis of some verb phrases. References will be given to Fig 2 which is an extract from text S.1.1 (TUs 222-269) in Svartvik & Quirk 1980:38. The symbols used for tags, surrounded by round brackets, are listed in Table 1.

Table 1: Tagging symbols for verbs

(V)	VERBS				
	MAIN VERBS (VA)	*BE* (VB)	*DO* (VD)	*HAVE* (VH)	*MODALS* (VM)
(+ϕ)	*speak*	*be*	*do*	*have*	
(+3)	*speaks*	*is*	*does*	*has*	
(+D)	*spoke*		*did*	*had*	
(+N)	*spoken*	*been*	*done*	*had*	
(+G)	*speaking*	*being*	*doing*	*having*	
(+1)		*am*			
(+4)		*are*			
(+5)		*was*			
(+6)		*were*			
(+8)					*can*
(+9)					*could*

CONTRACTIONS (indicated by asterisk), eg:

(VB+3*AN)	*isn't*
(VB+4*AN)	*aren't*
(RA*VB+4)	*you're*
(RA*VH+ϕ)	*I've*
(RA*VB+3)	*he's* (here)
(RA*VH+3)	*he's* (done it)

TU 222: *shouldn't* (VM+9*AN). Such a modal verb on its own, without a head, can be interpreted in several ways: a) a verb phrase fragment, as here (cf TU 223); b) an ellipted structure; c) a partial of a verb phrase which is continued in a following tone unit. A subsequent program will combine such partials in adjacent tone units into verb phrases.

TU 223: *shouldn't* (VM+9*AN) *be* (VB-ϕ) *imposing* (VA+G). This verb string satisfies our description of a complete verb phrase with the structure given by the automatic analysis as (VPH: modal progressive). The procedure is roughly as follows. First the rightmost (V) tag is spotted: (VA+G). Since it is an *-ing* form (+G), the program will search for any form of the verb *be* which progressive (complex) verb phrases are made up of. The search is successful since there is a previous form of *be* (VB) and the combination

(progressive) is noted. The next search is for some combination that fits the base form of *be* (VB+ϕ). It is successful since (VM) occurs left of (+ϕ) and the combination (modal) is noted. As a modal can only be the first element of a complex verb phrase, this search run for a complex verb phrase stops here. However, the scan proceeds to the left of (VM) in tone unit 223 in quest of another (V) which may constitute a new verb phrase. Since none is found, the verb phrase search for this tone unit stops and moves on to tone unit 224, starting from right to left, looking for a (V) . . .

TU 224: *'s* (VB+3). Since this verb phrase occurs in a stretch of transcription marked << >> (denoting subaudibility), it has not been analysed.

TU 225: also skipped because of subaudibility.

TU 226: *see* (VA+ϕ). Analysis: (VPH: present).

TU 223: *asking* (VA+G). Analysis: (VPH: ing), ie non-finite verb phrase consisting of present participle; which, in this case, will subsequently become part of a prepositional phrase.

TU 228: skipped because of subaudibility.

TU 229: skipped because no (V) in tone unit.

TU 230: *would* (VM+9) *be* (VB+ϕ). Analysis: (VPH: modal). Note that different functions of *be* are not marked in the word-class analysis, but all are (VB). One reason is of course that it is desirable to have a tagging system that is fully automatic on the word-class level, and functional differentation is therefore deferred until the phrase analysis stage. Unless *be* in TU 230 combines with another partial verb phrase in the text tone unit of speaker A, *be* in the verb phrase of TU 230 will be analysed as a copula.

TU 223: *'s* (VB+3) *ended* (VA+D,N). Analysis: (VPH: passive); *be* (VB+ϕ) *freed* (VA+D,N). Analysis: (VPH: passive). Both verb phrases are passive (but, for the rightmost, a better analysis would be for the program to propose alternatives (VPH: passive/perfective)). Failure to link the leftmost verb phrase *be freed* to the previous verb phrase contribution by speaker A, *would be* in TU 230 answers the query raised above. No complex verb phrase combination of (VB+ϕ) in TU 230 and (VB+ϕ) in TU 233 exists in

English grammar, hence *be* in TU 230 will be taken as a copula, not an aspectual auxiliary.

Figure 3 gives some examples of the output from a preliminary run of the verb phrase tagging program.

8 Summing up

The main purpose of the program outlined here is to give us information about the relation between grammar and prosody in large amounts of data representing different types of spoken English. The insights gathered from this analysis should help to answer some of the questions raised at the beginning of this paper by improving our understanding of how human speech is processed and thus enable us to produce better grammatical descriptions for spoken language.

The program can, in principle, be used also for grammatical analysis of written language and we are planning to make use of this facility in a future comparison between speech and writing.

We also hope that our work will make a contribution to the methodology of computational linguistics in that we will learn more about how far morphological syntactic analyses can be made automatic.

ACKNOWLEDGEMENTS

The Survey of Spoken English is supported by the Bank of Sweden Tercentenary Foundation. I want to acknowledge the contributions to the research reported here by the Survey staff: Mats Eeg-Olofsson (who is responsible for the programming), Oscar Forscheden, Bengt Orestrom and Cecilia Thavenius, as well as valuable comments by Bengt Altenberg on a draft version.

Figure 3: Examples of output from the verb phrase tagging program (Explanations of symbols used are given in Svartvik & Quirk 1980:21.)

```
TU 993  /you you /turn 'left and im:mediately RIGHT a'gain//
        VPH:present = turn(VA+0)

TU 884  there was an /old 'old WAY down 'onto that ROAD//
        VPH:past = was(VB+5)

TU 836   I should /think you'd A:LLOW two and a HALF//
        VPH:modal = should(VM+9) think(VA+0)
        VPH:modal = 'd(VM+9) allow(VA+0)

TU 957  /well !Frank's !practically never 'done it by :ROAD
        VPH:perfect = 's(VH+3) done(VD+N)

TU 989  /JO doesn't 'like 'this BIT//
        VPH:do-support = doesn't(VD+3) like(VA+0)

TU 994  in /order to 'go 'through !FARNHAM//
        VPH:infinitive = go(VA+0)

TU 627  /and then !digging TRENCHES//
        VPH:-ing form = digging (VA+G)

TU 528  /what WAS it 'called//
        VPH:passive = was(VB+5) called(VA+N)

TU 863  /if one's :coming up 'through :CROYDON//
        VPH:progressive = 's(VB+3) co ing(VA+G)

TU 927  /London would be worth going !THROUGH//
        VPH:modal = would(VM+9) be(VB+0)
        VPH:-ing form = going(VA+G)

TU 941  he's /been RE:SEARCHING it//
        VPH:perfective progressive = 's(VH+3) been(VB+N)
        researching VA+G)
TU 617  it /really would have been :much more DISASTROUS//
        VPH:modal perfective = would(VM+9) have(VH+0) been(VB+N)
```

[1] The London-Lund Corpus consists of 87 'texts' with origin in speech that have been recorded and transcribed at the Survey of English Usage, University College London and computerized at the Survey of Spoken English, University of Lund. For a description of the corpus, see Svartvik & Quirk (1980).

[2] One feature related to planning is speed of production. It is estimated that the average speed of spoken English is about 180 words per minute and that hand-writing is slower than one-tenth the speed of speaking (Chafe, forthcoming: 4).

[3] However, some psychologists are beginning to realize the value of large data bases: "there is a problem ... of securing a corpus of data that is truly adequate to test conclusions decisively", "Clearly, the ultimate solution ... must be to automate the processing of long streams of speech, so as to detect ungrammaticality automatically, and thereby enlarge the sample of syntactic errors to hundreds or thousands of examples instead of a handful ... The syntactic analysis would not have to be exact – it must simply beneficiate the raw ore so that human experts could afford to search through the enriched samples" (Simon 1980: 543-544).

[4] Lengthening towards the end of the tone unit may also have a bearing on theme/rheme organization within the tone unit (see, for example, Kroch & Hindle 1981), but this point will not be further pursued here.

[5] For a definition of 'tone unit', see Crystal 1969:204ff. Linguists generally agree on the identification of tone units. Quirk & Crystal (1966) also report 84.8 per cent agreement on tone unit identification among 46 informants untrained in prosodic analysis. However, Brown *et al.* report difficulty with consistently identifying tone groups by intonational criteria but recognize 'pause-defined units' (1980:20). Chafe uses three criteria: intonation, pausing, and syntax (1980:14).

[6] Cf McNeill: 'the unit of speech programming is not a grammatical constituent in many utterances. This is not without theoretical interest. The performance device in these utterances cannot be said to have more than *weak equivalence* with the grammatical description ... and not, as widely assumed ... a strong equivalence ...' (1979:250).

REFERENCES

Aaronson, D. & R.W. Rieber (eds.). 1979. Psycholinguistic research: implications and applications. Hillsdale, New Jersey: Lawrence Erlbaum.

Brown, G., K.L. Currie & J. Kenworthy. 1980. Questions of intonation. London: Croom Helm.

Brown, K. 1980. Grammatical incoherence. In Dechert & Raupach 1980, 27-37.

Chafe, W.L. 1980. The deployment of consciousness in the production of a narrative. The pear stories, ed. by W.L. Chafe, 9-50. Norwood, New Jersey: Ablex.

— Forthcoming. Integration and involvement in speaking, writing, and oral literature. Spoken and written language, ed. by D. Tannen. Norwood: N.J.: Ablex.

Clark, H.H. & E.V. Clark. 1977. Psychology and language. An introduction to psycholinguistics. New York: Harcourt Brace Jovanovich.

Cole, R.A. (ed.). 1980. Perception and production of fluent speech. Hillsdale, New Jersey: Lawrence Erlbaum.

— & J. Jakimik. 1980. A model of speech perception. In Cole 1980, 133-163.

Cooper, W.E. & J. Paccia-Cooper. 1980. Syntax and speech. Cambridge, Mass.: Harvard University Press.

Crystal, D. 1969. Prosodic systems and intonation in English. Cambridge: Cambridge University Press.

— 1980. Neglected grammatical factors in conversational English. In Greenbaum *et al.* 1980, 153-166.

Dechert, H.W. & M. Raupach (eds.). 1980. Temporal variables in speech. Studies in honour of Frieda Goldman-Eisler. The Hague: Mouton.

Greenbaum, S., G. Leech & J. Svartvik (eds.). 1980. Studies in English linguistics for Randolph Quirk. London: Longman.

Halliday, M.A.K. 1967. Notes on transitivity and theme in English. Journal of Linguistics 3:1, 37-81.

Kroch, A.S. &. D. Hindle. 1980. A quantitative study of the syntax of speech and writing. Progress report to the N.I.E.

Lieberman, P. 1980. The innate, central aspect of intonation. In Waugh and van Schooneveld 1980, 187-199.

McNeill, D. 1979. Natural processing units of speech. In Aaronson & Rieber 1979, 241-261.

Quirk, R. & D. Crystal. 1966. On scales of contrast in connected English speech. In memory of J.R. Firth, ed. by C.E. Bazell, H.C. Catford, M.A.K. Halliday & R.H. Robins, 359-369. London: Longman.

Quirk, R., J. Svartvik, A.P. Duckworth, J.P.L. Rusiecki & A.J.T. Colin. 1964. Studies in the correspondence of prosodic to grammatical features in English. Proceedings of Ninth International Congress of Linguistics, 679-691. The Hague: Mouton.

Simon, H.A. 1980. How to win at twenty questions with nature. In Cole 1980, 535-548.

Sternberg, S., C.E. Wright, R.L. Knoll & S. Monsell. 1980. Motor programs in rapid speech: additional evidence. In Cole 1980, 507-534.

Svartvik, J. & M. Eeg-Olofsson. Forthcoming. Tagging the London-Lund Corpus of spoken English. Computer corpora in English language research, ed. by Stig Johansson. Report from the Norwegian Computing Centre for the Humanities, Bergen.

Svartvik, J. & R. Quirk (eds.). 1980. A corpus of English conversation. Lund Studies in English 58. Lund: Liber/Gleerup.

Waugh, L.R. & C.H. van Schooneveld (eds.). 1980. The melody of language. Baltimore: University Park Press.

Weiner, S.L. & D.R. Goodenough. 1977. A move towards a psychology of conversation. Discourse processes: advances in research and theory 1: Discourse production and comprehension, ed. by R.E. Freedle. Norwood, New Jersey: Ablex.